8-1-2003

The Perception and Recognition of 3-D Shape from Shadows Cast onto Curved Surfaces

Young-Lim Lee
Western Kentucky University

THE PERCEPTION AND RECOGNITION OF 3-D SHAPE FROM SHADOWS CAST ONTO CURVED SURFACES

A Thesis

Presented to

the Faculty of the Department of Psychology

Western Kentucky University

Bowling Green, Kentucky

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

By

Young-Lim Lee

August 2003

THE PERCEPTION AND RECOGNITION OF 3-D SHAPE FROM

SHADOWS CAST ONTO CURVED SURFACES

Date Recommended 6/2/2003

J. Farley Norman

Director of Thesis

Elmer Gray 7/14/03

Dean, Graduate Studies and Research Date

Acknowledgments

First of all, I would like to thank the director of this thesis, Dr. J. Farley Norman, for his direction, support, and dedication throughout my study at Western. Dr. Norman is the greatest advisor and mentor, and he and his wife, Dr. Hideko Norman, are like my family. I would also like to thank the other members of my committee, Dr. Joseph Bilotta and Dr. Dan Roenker, for their suggestions and support. I wish to thank my family and friends both here and in my country, Korea, for their love and encouragement. Lastly, I have to thank God for guiding and blessing me.

Table of Contents

List of Figures

THE PERCEPTION AND RECOGNITION OF 3-D SHAPE FROM

SHADOWS CAST ONTO CURVED SURFACES

Young-Lim Lee August 2003 27 Pages

Directed by: J. Farley Norman, Ph.D., Joseph Bilotta, Ph.D., and Dan Roenker, Ph.D.

Department of Psychology Western Kentucky University

The informativeness of a cast shadow or silhouette boundary contour for the perception of 3-D object shape has been investigated for a long time. Some researchers have focused on the informativeness of static shadows (Attneave, 1954; Koenderink, 1984; Richards, Koenderink, & Hoffman, 1987; Norman, Phillips, & Ross, 2001; Tse, 2002) while other researchers have focused on the informativeness of moving or deforming shadows (Miles, 1931; Wallach, & O'Connell, 1953; Norman, & Todd, 1994; Norman, Dawson, & Raines, 2000; Norman, & Raines, 2002). Past research has shown that changing the angle of illumination does not affect the perception of 3-D shape from cast shadows (Norman et al., 2000). The current experiment extends the prior experiments by further investigating whether curved background surfaces (with both positive and negative Gaussian curvature) affect the perception and recognition of 3-D object shape from deforming and/or static cast shadows. In this experiment, the observers viewed either deforming or static shadows of naturally shaped objects (bell-peppers) cast onto either flat, hemispherical, or saddle-shaped surfaces. The results revealed significant main effects of motion (deforming vs. static shadows), object, and the type of background surface. The results revealed that there were also a number of significant interactions involving particular objects, the presence or absence of motion, and the type of background surface. The observers' ability to recognize objects from deforming shadows was higher than their ability to recognize objects from static shadows. In addition, the observers' ability to recognize objects from the shadows cast onto the hemisphere background surface was generally as accurate as their ability to recognize objects from the shadows cast onto the flat plane. However, the observers' ability to recognize the objects was reduced when the shadows were cast onto the saddle background surface. The results of the experiment confirm previous findings showing that shadow boundary contours, especially deforming contours, are perceptually informative and help observers to perceive and recognize 3-D object shape. This experiment also extends previous studies by showing how differently curved background surfaces affect the perception and recognition of 3-D object shape.

Chapter 1

Introduction

Human observers can obtain useful information about the three-dimensional (3-D) shape of objects from their silhouettes or cast shadows. This issue has long been studied by many investigators. For example, Attneave (1954), Koenderink (1984), Richards, Koenderink, and Hoffman (1987), Norman, Phillips, and Ross (2001), and Tse (2002) have all studied the importance of static shadows for perceiving the shape of 3-D objects. In addition, Miles (1931), Wallach and O'Connell (1953), Norman and Todd (1994), Norman, Dawson, and Raines (2000), and Norman and Raines (2002) showed that human observers could more accurately perceive the shape of 3-D objects from moving shadows than from static shadows.

In the case of static shadows, Koenderink (1984) emphasized the importance of their outer boundary contours for perceiving the shape of 3-D objects. He suggested, based upon mathematics, that the most informative parts of a shadow or silhouette boundary contour are its convexities and concavities. He said that "some silhouettes are ambiguous and some are not" (p. 327). Ambiguous shadows or silhouettes usually lack prominent convexities and concavities. In other words, some silhouettes of solid objects look solid and 3-dimensional, but some look flat. This ambiguity may occur when viewing an object from an unusual vantage point. For example, in a head-on or direct view, the silhouette of a cylinder is ambiguous (i.e., it has the same projected shape as a rectangle). However, a slightly different viewpoint will often lead to a more recognizable silhouette. As Figure 1 shows, silhouettes of a single 3-D object can be quite different depending on the observers' viewpoint (i.e., Figure 1 illustrates two views of a

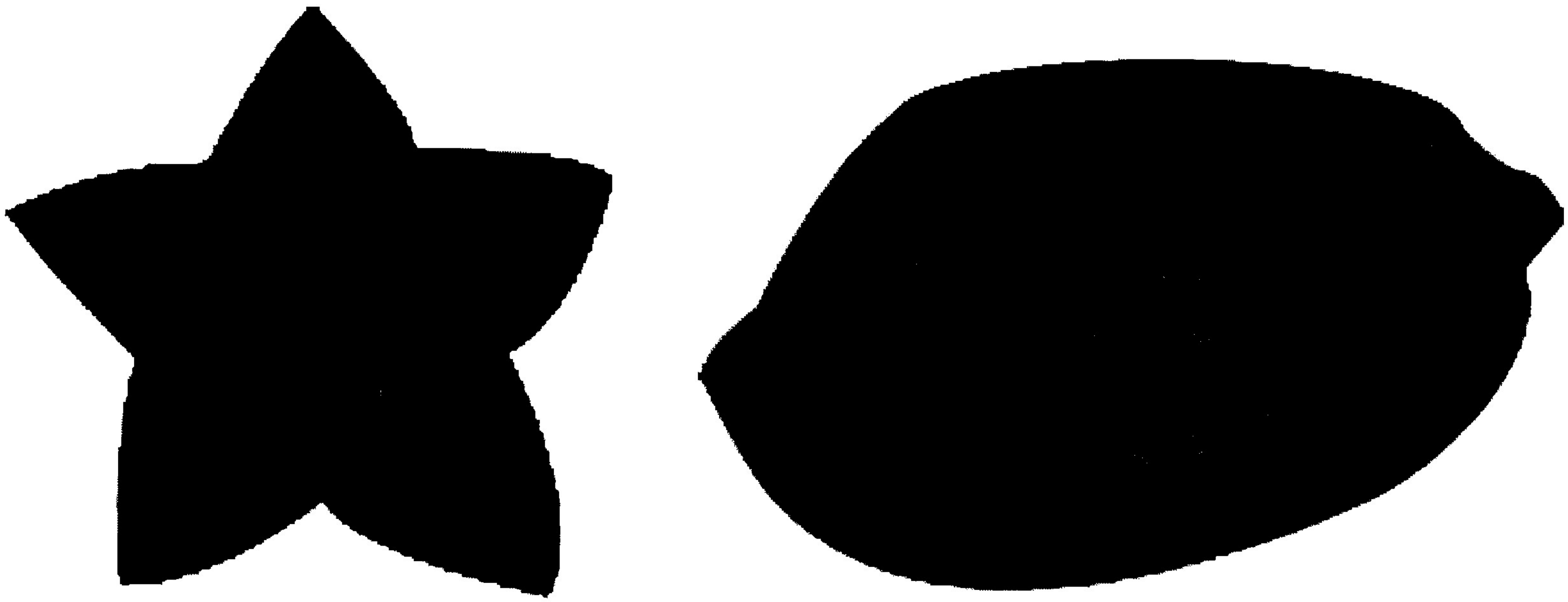


Figure 1. Two different silhouettes of a starfruit. Notice that the silhouettes are quite different, depending upon the orientation of the object or viewpoint of an observer.

"starfruit," *Averrhoa Carambola* - a star-shaped silhouette is seen from one vantage point, while an elliptical silhouette is seen from a different viewpoint). Therefore, it is sometimes difficult to recognize the 3-D shape of an object from static 2-D silhouettes.

Richards et al. (1987) also showed how the interpretation of the shape of 3-D objects could be derived from 2-D silhouettes. They demonstrated that there is a relationship between the curvature of various regions on the surface of 3-D objects and the type of curvature present within 2-D silhouettes. In particular, convexities in a 2-D shadow or silhouette indicate that there is a "bump" on the surface of the 3-D object casting the shadow, while concavities in the 2-D silhouette indicate that the corresponding surface region on the 3-D object is locally shaped like a saddle. Richards et al. also recommended that in inferring 3-D shape from 2-D silhouettes that if there are no undulations in the silhouette, then one should not postulate undulations (i.e., bumps or saddles) on the surface of the 3-D object. However, this recommendation is not always appropriate, because (as shown in Figure 1) the outline of 2-D silhouettes or shadows can sometimes be unusually simple and not reflect the true shape of the 3-D object, especially if the object is viewed from an unusual position. Therefore, human observers cannot always perceive the shape of 3-D objects correctly when viewing static 2-D silhouettes. This limitation is one of the reasons why researchers have been interested in the informativeness of moving shadows for the perception of 3-D shape.

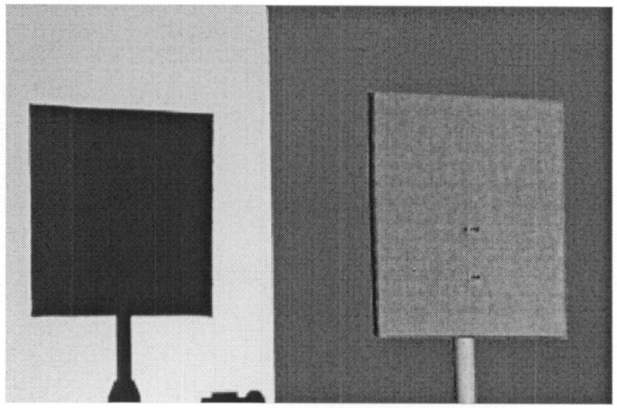
In 1953, Wallach and O'Connell used moving 3-D objects such as wooden blocks, cylinders, wire-frame figures, and rods and cast their shadows onto a projection screen. They found that when human observers viewed the deforming 2-D silhouettes, they perceived solid 3-D objects rotating in depth. They referred to this phenomenon as the 'kinetic depth effect.' However, Wallach and O'Connell also found that the kinetic depth effect occurred only when identifiable regions such as the endings of linear boundary contours or sharp corners were present in the deforming silhouettes. For example, they concluded that "curved contours which are deformed without displaying a form feature which identifies a specific point along the curve are seen as distorting, often even if for some reason the shadow is seen as a 3-D form … it is now clear that the perceived distortions of deforming curved contours are not related to the kinetic depth effect at all" (p. 209).

More recent research has shown that Wallach and O'Connell's conclusions regarding the deformation of shadows with smoothly curved boundaries were not entirely correct. Norman and Todd (1994) found that the deformation of smoothly curved boundary contours can lead to the perception of rigidly moving 3-D objects. They used ellipsoids whose 2-D silhouettes did not have sharp corners or specific identifiable regions. However, the study of Norman and Todd (1994) has limited generalizability because the shape of ellipsoids is much simpler than that of most natural objects (such as the starfruit depicted in Figure 1).

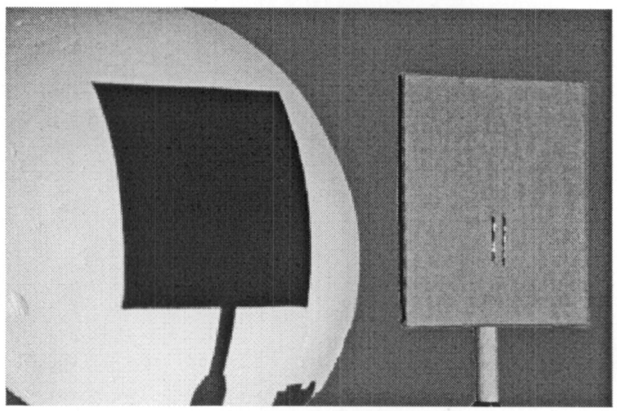
In an attempt to further study the informativeness of shadows for the perception of 3-D object shape, Norman et al. (2000) used more naturally shaped objects (bell peppers). Their results showed that observers were able to recognize specific 3-D objects from the deformations of their projected cast shadows, even though the different bell peppers were very similar in shape. Their results also showed that there were no effects of variations in the angle of illumination despite the fact that changes in the angle of illumination had large effects upon the specific shape of the cast shadows. The finding that observers could perceive 3-D shape and rigid rotations in depth from such smoothly curved shadow deformations confirmed the previous study of Norman and Todd (1994).

Norman et al. (2001) specifically examined Attneave's (1954) hypothesis that the most informative parts of a silhouette or shadow boundary contour are points of maximal curvature. They investigated the informativeness of the boundary contours of naturally shaped 3-D objects (sweet potatoes). They asked the observers to look at the silhouettes of objects and draw or copy the shape of the silhouettes. For each object, the observers adjusted the positions of ten dots until the dotted figure had the same shape as the original silhouette. After finishing their dotted drawing, the observers were asked to place marks on the original silhouette boundary contour corresponding to each dot's final adjusted position. Norman et al. calculated the curvatures along the silhouette boundaries to test Attneave's hypothesis and found that observers did choose regions with locally high curvatures. Therefore, their results were consistent with Attneave's (1954) conclusion that curvature maxima are the most informative parts of boundary contours for the human perception of object shape.

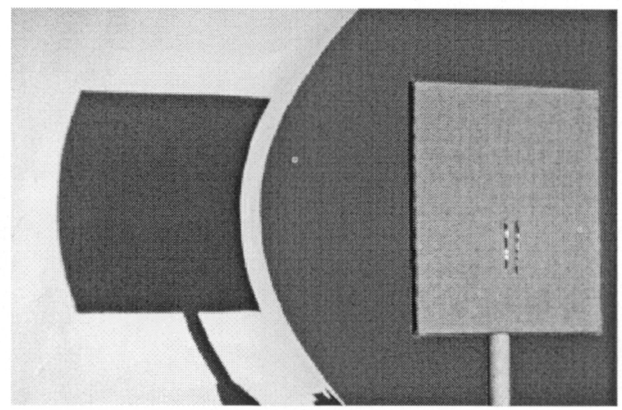
The current experiment is an extension of previous studies that have investigated how observers perceive 3-D objects from their cast shadows. All previous researchers used flat projection screens upon which the shadows of 3-D objects were cast. However, in a natural environment, shadows are usually cast onto curved object surfaces such as rocks, wooden tree trunks, etc. The shape of cast shadows is as much affected by the nature of the background surface as by the shape of the object itself. This phenomenon can be seen in Figure 2, which shows the shadows of a square cast onto flat and curved background surfaces. Notice that the straight edges of the square project to curved edges



Flat Plane



Hemisphere



Saddle

Figure 2. Photographs of shadows of a square cast onto flat and curved background surfaces.

in the shadows cast onto curved background surfaces. In this experiment, the effects of different background surfaces were investigated. In particular, the observers in this experiment were required to recognize shadows cast onto surfaces with both positive (e.g., hemispheres) and negative Gaussian curvature (e.g., saddles).

Chapter 2

Method

Observers

The stimulus displays were presented to four observers (HFN, MCW, RM, XW) who were either students or faculty at Western Kentucky University. Half of the observers were males and half were females. All observers had normal or corrected-to-normal visual acuity.

Apparatus

An Apple dual-processor G4 Power Macintosh was used to generate and display the cast shadows of the objects. The stimulus patterns were displayed on a Mitsubishi Diamond Plus 200 22-inch monitor with a 1280x1024 pixel resolution. The stimulus displays were accelerated using a Radeon 8500 graphics accelerator card (ATI Technologies, Inc). The observers viewed the stimulus displays at a viewing distance of 100 cm from the computer monitor.

Stimulus Displays

The objects used to generate the shadows were plastic replicas of five ordinary bell peppers (made out of Smooth-Cast 321 liquid plastic, Smooth-on, Inc.). The five peppers were chosen to have similar sizes in order to prevent recognition based on overall differences in size. Shadows of these five objects were cast onto either a flat translucent screen or upon hemispherical or saddle-like curved surfaces (radius of curvature = 14 cm). A Nikon Coolpix 995 digital camera was used to record the cast shadows, and the shadow images were then transferred to the computer. In order to obtain the clearest and sharpest shadows, we placed the objects 20 cm in front of the background surface. The distance from the camera to the objects was approximately 110 cm. The light source used to cast the shadows was a 300 watt tungsten halogen lamp (Sylvania ELH), and an additional hood was placed around the lamp to direct the light towards the object and background surface. The distance between the light source and the objects was 6.5 m, and the angle between the camera's line of sight and the direction of the light source was 35 degrees. Eighty different shadows of each object were obtained by rotating the objects 360 degrees in depth around a vertical axis in 4.5 degree angular increments. Sample shadows of each of the five objects (cast onto a flat background surface) are presented in Figure 3.

Procedure

All observers participated in a series of practice trials with feedback and experimental trials without feedback. Practice trials were needed for the observers to learn to make an appropriate response to the five objects prior to the start of an experimental session. In the practice trials, the observers were required to identify an object on every trial. Consecutive series of ten trials (deforming shadows of 5 objects cast onto the flat background surface x 2 practice trials per object) were presented until each observer's recognition accuracy reached 90 percent or higher. During the practice trials, the observers received feedback in the form of a short auditory beep for a correct answer. Once this 90 percent criterion was reached, the experimental session began.

Each of the observers participated in ten sessions. Within each session, the observers were presented with a single block of 300 trials (30 experimental conditions x 10 trials per condition). The 30 conditions consisted of 5 objects x 3 background surfaces

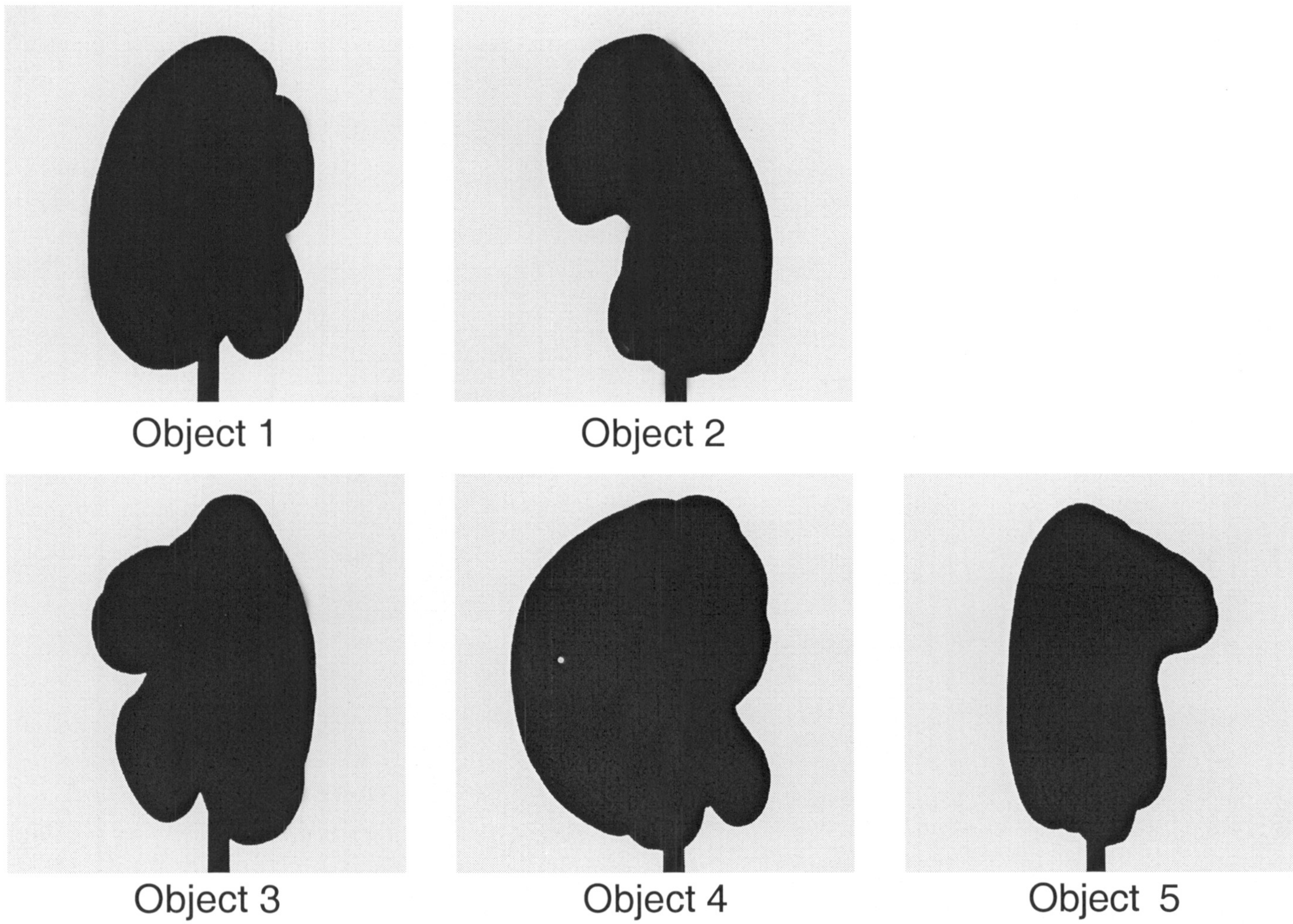


Figure 3. Representative shadows for each of the five stimulus objects.

(flat, hemisphere, saddle) x 2 motion types (deforming shadows versus static shadows). The order of the conditions within a block was randomly determined for each observer. For a condition depicting a static shadow, one of the 80 captured shadows for an object was chosen randomly. In conditions depicting deforming shadows, full rotations of the objects (i.e., all 80 shadows) were presented. The observers were required to identify an object (1-5) from either type of display (deforming or static shadows), and then press an appropriate key on the computer's keyboard. The observers had as much time as they needed to view each shadow and identify the object. The observers never received feedback during an experimental session. After completion of all ten experimental sessions, a total of 100 responses had been collected for each of the 30 conditions (3000 total trials per observer).

Chapter 3

Results

The results for all observers (HFN, MCW, RM, XW) are shown in Figures 4-7, in which recognition accuracy is plotted as a function of object, motion type (deforming vs. static), and type of background surface. There were significant main effects involving motion type, objects, and type of background surface. There was a large main effect of motion (within-subjects factorial ANOVA, $F(1,3) = 130.7$, $p = .0014$). The observers' performance was nearly perfect for recognition of objects defined by deforming shadows (except for some object shadows projected onto the saddle) but was significantly reduced for static shadows. The mean recognition accuracy for all observers was 90 percent correct for the deforming shadows and 61 percent correct for the static shadows. In addition, there was a main effect of objects ($F(4,12) = 4.8$, $p = .015$). Across all observers, the mean recognition accuracy for each object was 72, 75, 67, 74, and 92 percent correct for objects 1-5, respectively. The observers' recognition performance for object 5 was higher than that observed for objects 1-4. Therefore, object 5 was the easiest to recognize. There was also a large effect of the type of background surface ($F(2,6) = 30.5$, $p = .0007$) such that performance was high for the recognition of objects cast onto the flat and hemisphere background surfaces but was significantly reduced when object shadows were cast onto saddle-shaped background surfaces. Across all observers, the mean recognition accuracy was 83 percent correct for shadows cast onto the flat surface, 80 percent correct for shadows cast onto the hemisphere, and 65 percent correct for shadows cast onto the saddle background surface. It is important to remember that chance recognition performance for this task would be 20 percent correct.

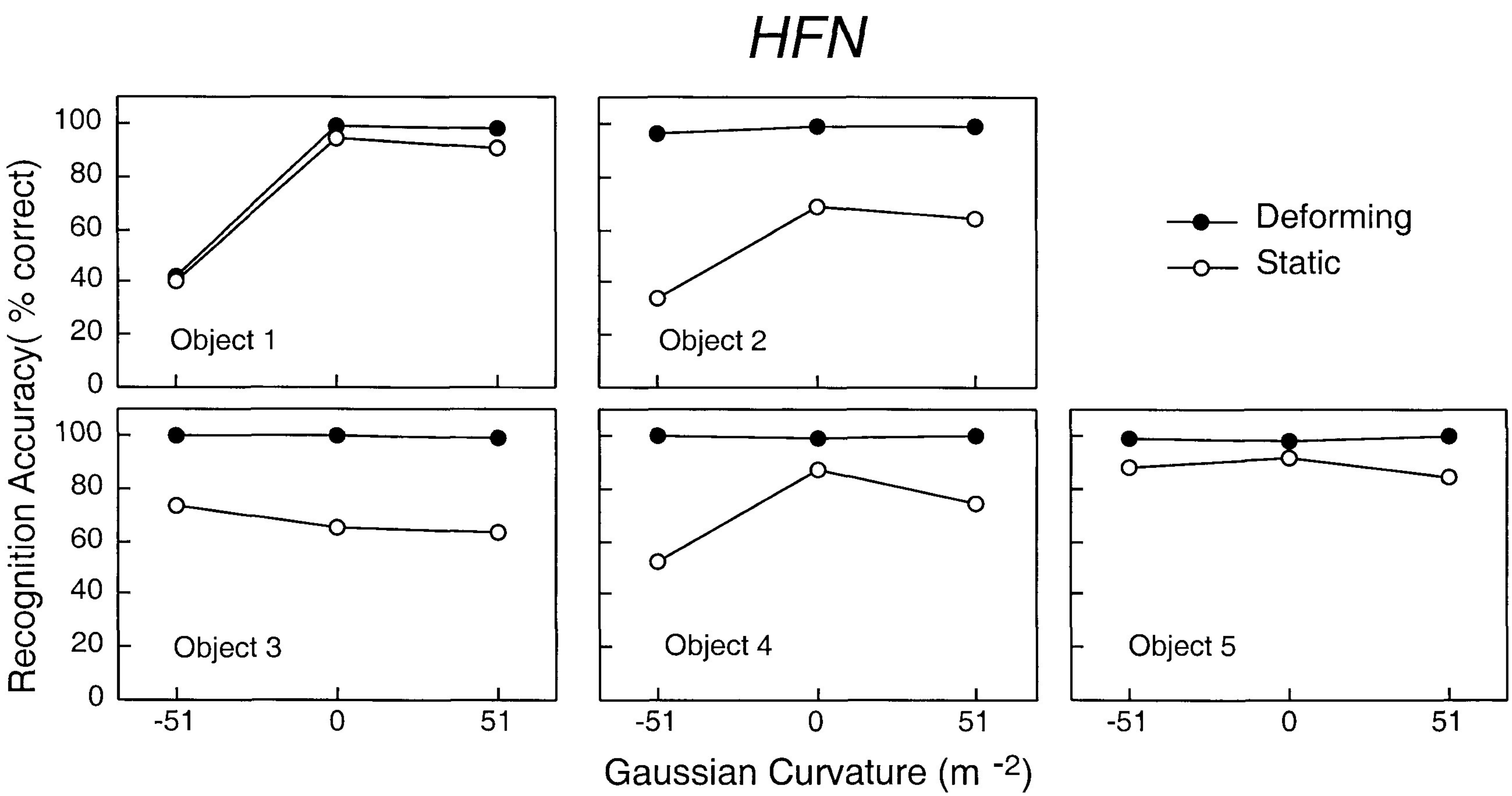


Figure 4. Results of the experiment for observer HFN. Saddle-shaped surfaces possess negative Gaussian curvature, while hemispherical surfaces possess positive Gaussian curvature.

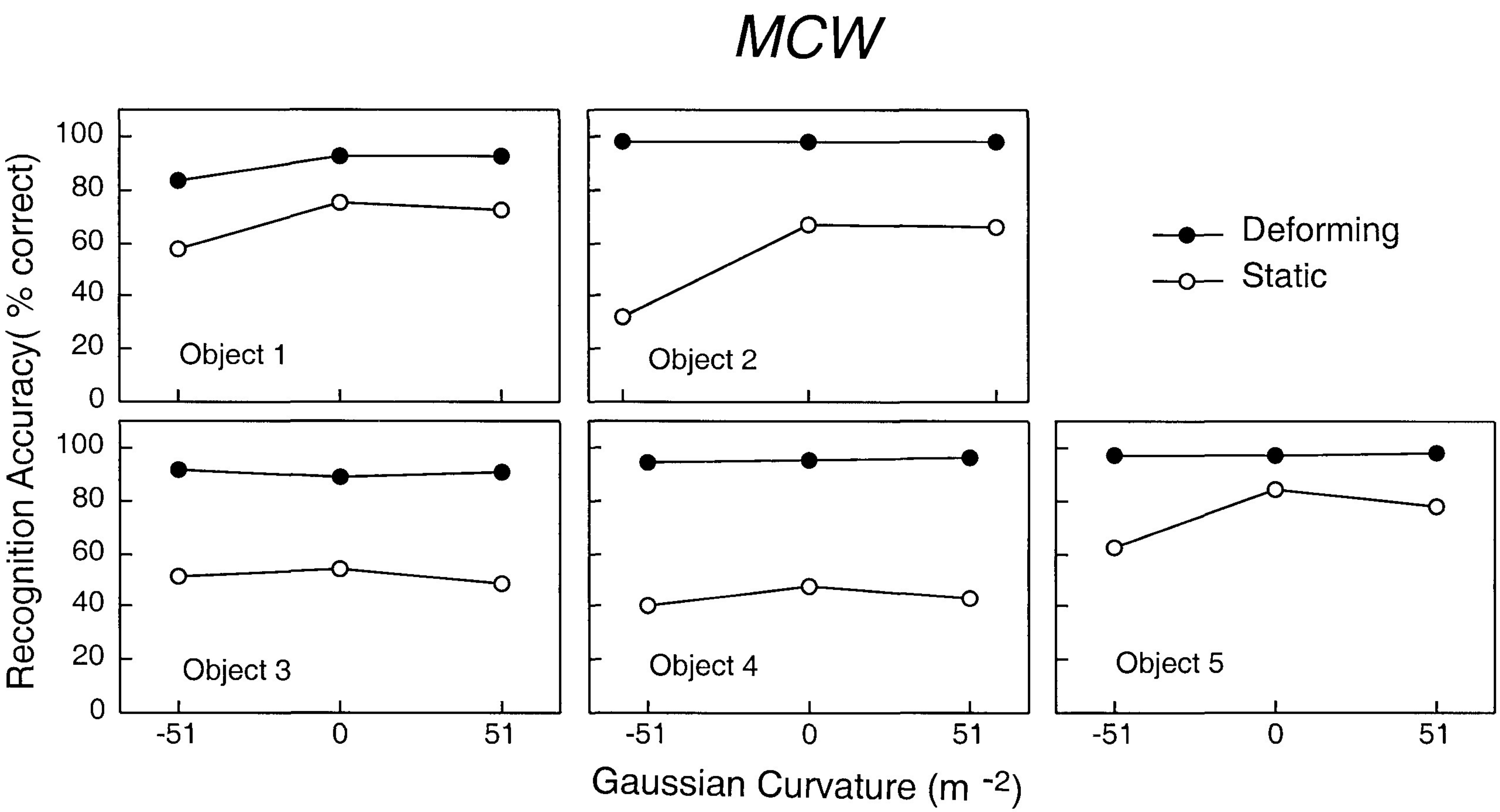


Figure 5. Results of the experiment for observer MCW. Saddle-shaped surfaces possess negative Gaussian curvature, while hemispherical surfaces possess positive Gaussian curvature.

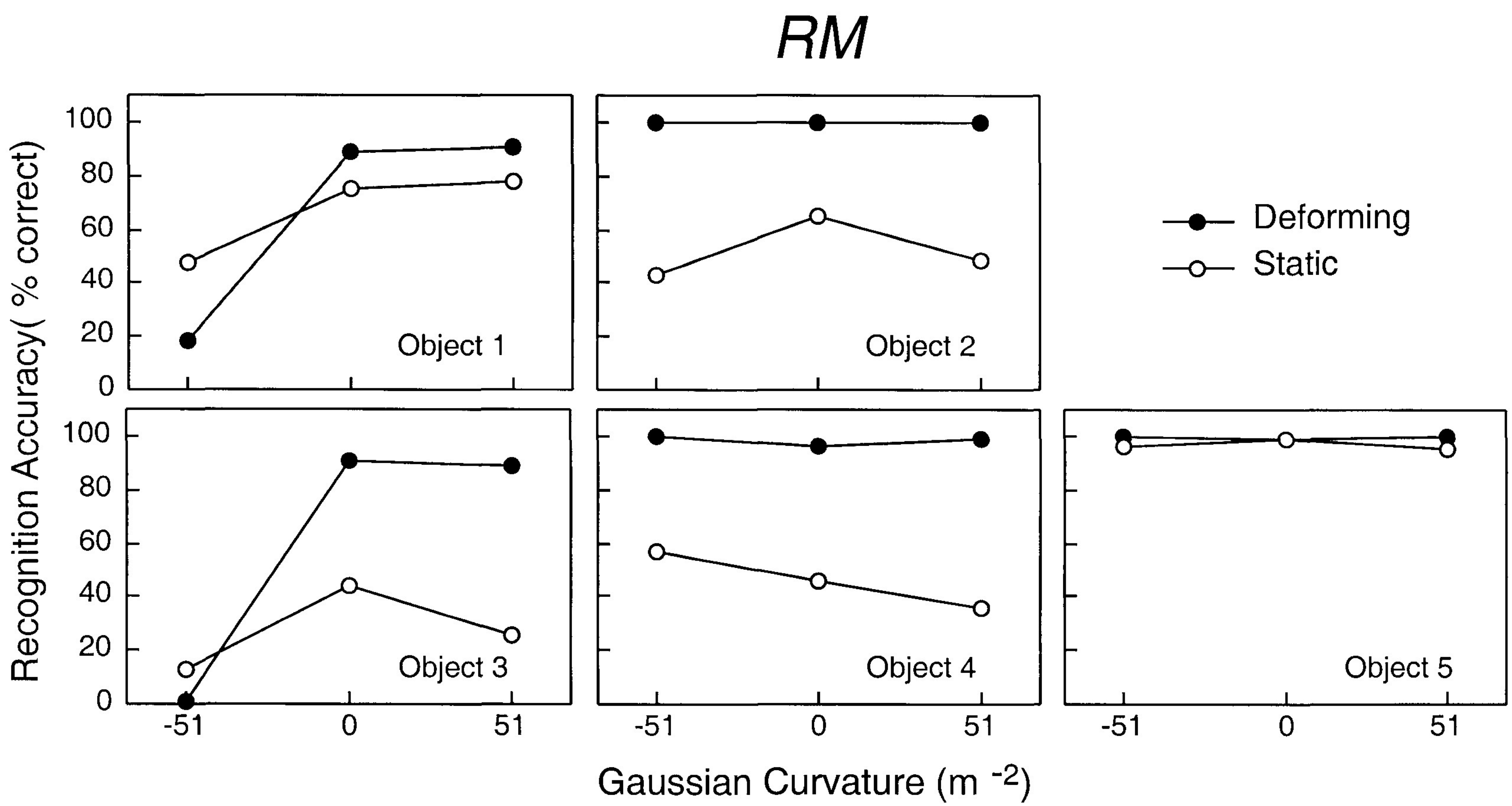


Figure 6. Results of the experiment for observer RM. Saddle-shaped surfaces possess negative Gaussian curvature, while hemispherical surfaces possess positive Gaussian curvature.

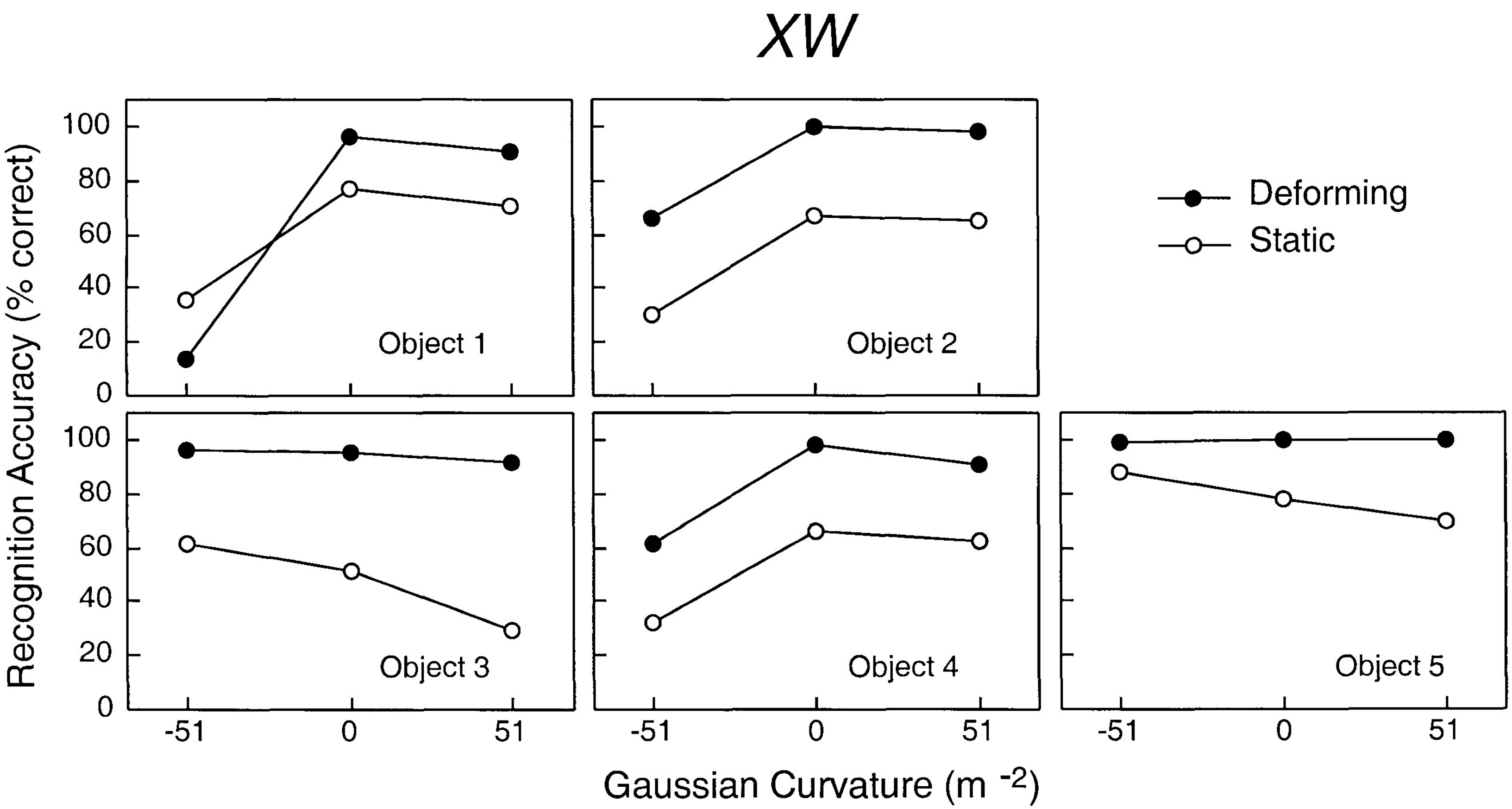


Figure 7. Results of the experiment for observer XW. Saddle-shaped surfaces possess negative Gaussian curvature, while hemispherical surfaces possess positive Gaussian curvature.

It is evident, therefore, that the observers were able to recognize the objects correctly in most instances.

Other significant interaction effects are presented in Figures 8 and 9. There was a large motion x object interaction ($F(4,12) = 13.9$, $p = .0002$). The absence of motion in the static shadows negatively affected the recognition of some objects (e.g., objects 2, 3, and 4, see Figure 8) more than others (e.g., objects 1 and 5). In addition, there was a large motion x object x background surface interaction ($F(8,24) = 5.8$, $p = .0003$). Across all observers and conditions, there was the same pattern of results for both the flat and hemisphere background surfaces, while the pattern of results for the saddle background surface was quite different (see Figure 8). The observers' performance for the deforming shadows was nearly perfect for all objects cast onto the flat and hemisphere background surfaces but was reduced for the shadows of objects 1 and 3 cast onto the saddle background surface. There was also a significant object x background surface interaction ($F(8,24) = 2.5$, $p = .037$), shown in Figure 9. One can see that the recognition performance was worst for shadows cast onto the saddle, particularly for objects 1 and 2. In contrast, objects 1 and 2 were relatively easy to identify from the shadows that were cast onto the hemisphere or flat plane.

The observers' recognition performances for the static shadows were analyzed in more detail. Consider, for example, Figure 10. Observers HFN and XW confused the shadows of object 2 that were cast onto the saddle with those of object 5. This confusion is understandable, since the shadows of object 2 cast onto the saddle are very similar to those of object 5 cast onto the plane (for example, compare the right two panels of Figure 10). In fact, there is as much resemblance between the shadows of object 5 and object 2

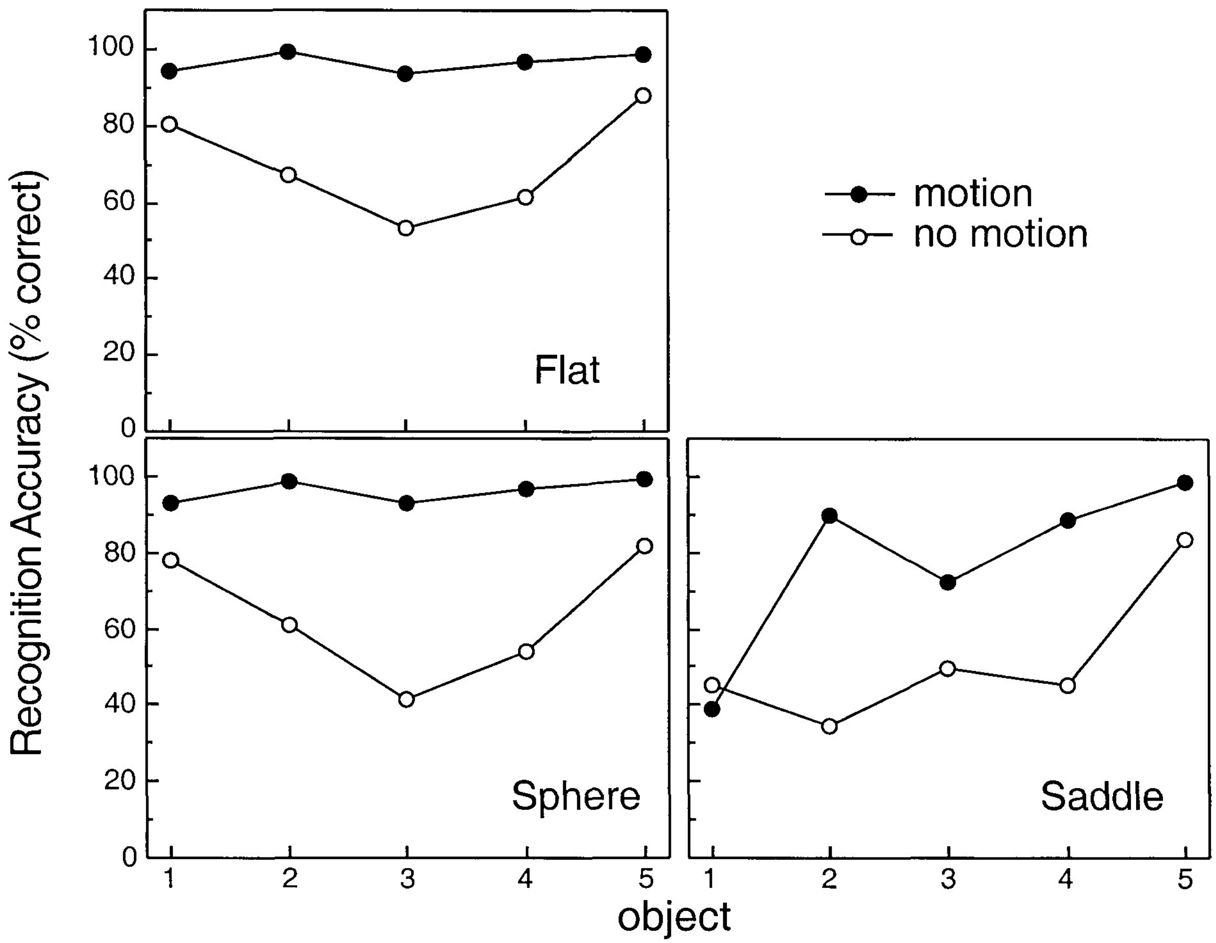


Figure 8. Experimental results for all observers, showing effects of motion, object, and type of background surface.

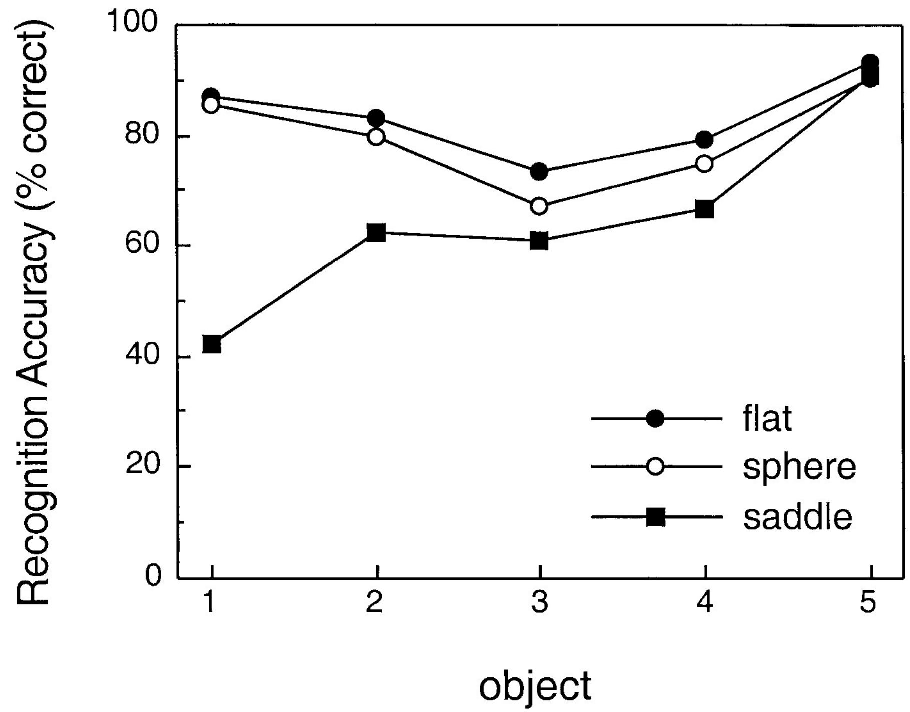


Figure 9. Experimental results for all observers showing the interaction involving the objects and type of background surface.

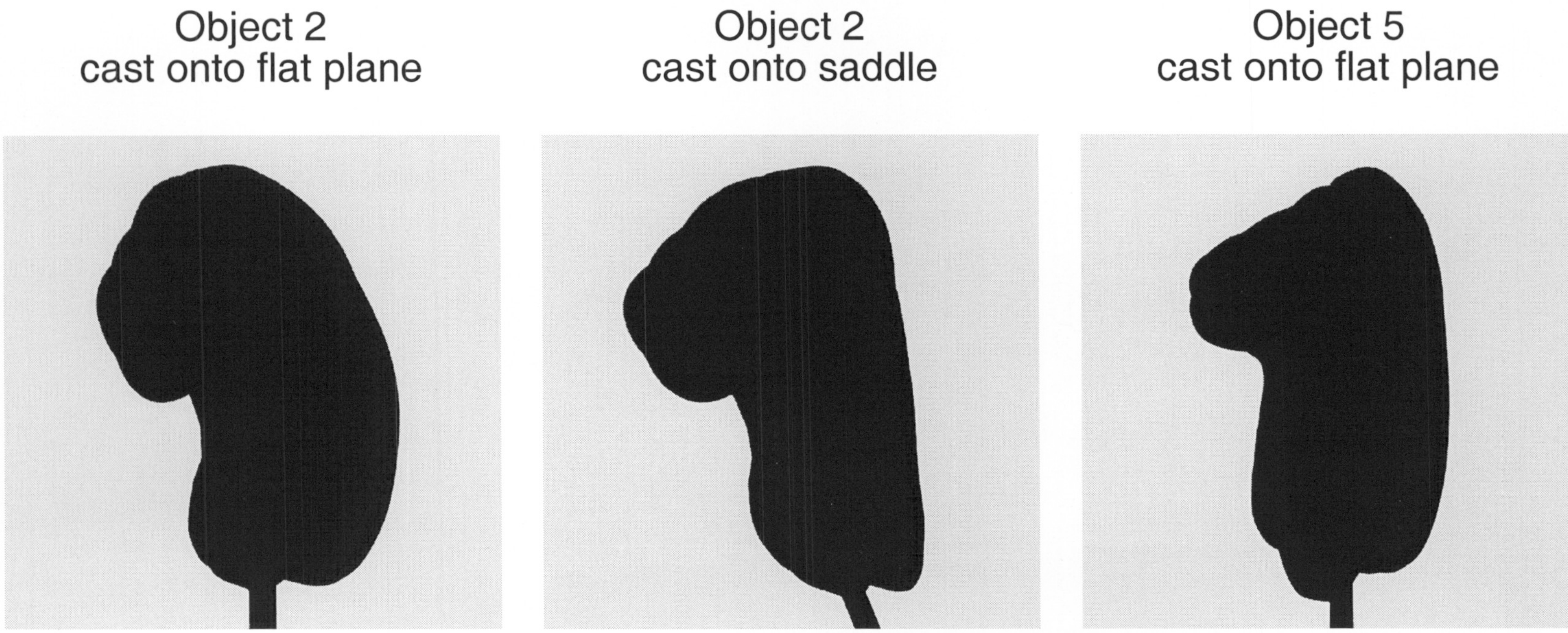


Figure 10. Shadows of objects 2 and 5. Notice that the shadow of object 2 cast onto the saddle closely resembles that of object 5 cast onto the flat plane.

(right two panels) as between object 2 with itself (left two panels). In order to examine such object confusions in more detail, a confusion analysis was conducted. Confusion matrices were obtained for all pairs of objects for each observer, and $- \ln \eta$ values were calculated (Luce, 1963). A $\ln \eta$ value of 0.0 indicates that two stimuli are not discriminable (i.e., are confusable), while values of 4.0 or higher indicate that two stimuli are highly discriminable (i.e., not confusable). The number of observers who confused (i.e., had low discrimination performance indicated by values of $- \ln \eta$ below 2.0) each of the ten pairs of objects for each type of background surface are shown in Figure 11. Some pairs (e.g., objects 1 and 5, 3 and 5, and 4 and 5) were never confused for any type of background surface, while other pairs (e.g., objects 1 and 3, 1 and 4, 3 and 4, and 2 and 5, see Figure 11) were confused by nearly all of the observers. The static shadows of objects 2 and 5 were confused by only two observers when cast onto the flat background surface, while all observers confused object 2 with object 5 and vice versa for the hemisphere and saddle background surfaces. In a similar vein, none of the observers confused the shadows of objects 2 and 4 when cast onto either the flat plane or hemisphere, but fully half of the observers did so when they were cast onto the saddle. One can see from the data shown in Figure 11 that the observers exhibited 40 percent more object pair confusions when the shadows were projected onto the saddle as compared to the hemisphere.

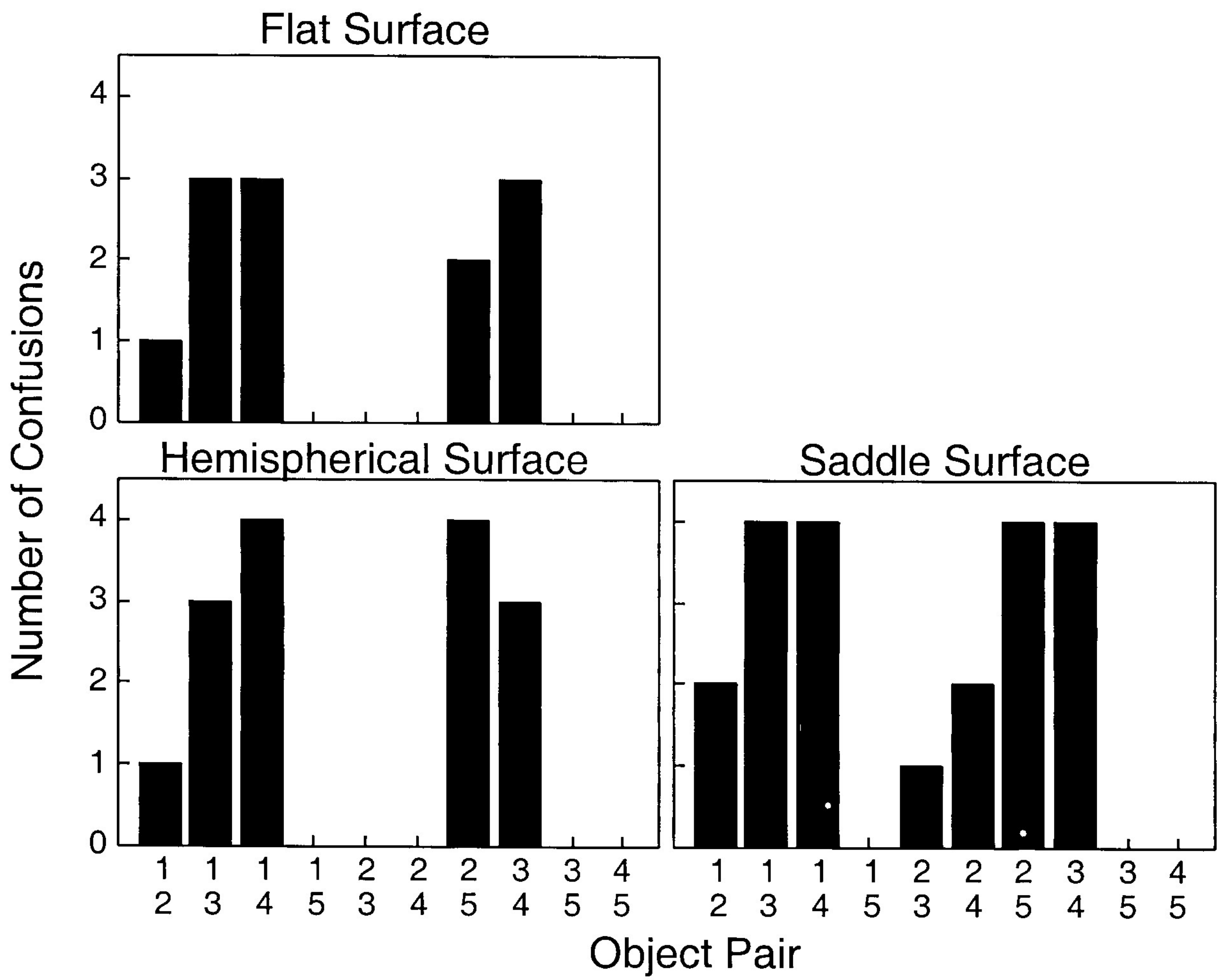


Figure 11. Results of the confusion analysis for all observers.

Chapter 4

Discussion

The results of the current experiment extend previous investigations that have examined the informativeness of cast shadows and boundary contours for the perception of 3-D object shape. This experiment examined how the presence of curved background surfaces affected the observers' ability to recognize 3-D objects from both deforming and static cast shadows. Previously, Norman et al. (2000) found that the perception of 3-D object shape from cast shadows was invariant over changes in the 2-D shape of the shadows that were induced by movements of the light source. The current experiment evaluated whether the perception of 3-D object shape is similarly invariant over changes in the 3-D shape of the background surface. The results showed that the observers' recognition performances from shadows cast onto the flat and hemisphere background surfaces were similar, but their performances from shadows cast onto the saddle background surface were quite different (see Figures 8 and 9). The observers also exhibited more confusion between pairs of objects from shadows cast onto the saddle background surfaces (see Figure 11). The saddle background surface negatively affected the perception and recognition of 3-D shape from both the deforming and static shadows. However, it is important to note that the deterioration of performance for shadows cast onto the saddle does not occur simply because casting shadows onto curved surfaces distorts the projected 2-D shape of the shadows (see Figure 2), since hemispherical background surfaces also distort 2-D projected shape. In general, the observers accurately recognized the 3-D objects from their shadows that were cast onto the hemisphere. Indeed, their performance for these conditions was almost identical to that obtained from shadows cast onto the flat plane. It is not clear as to why the perception of 3-D object shape from shadows was worse when the shadows were cast onto the saddle background surface. However, it might have occurred because hemispherical objects can be easily found (e.g., large round rocks, etc.) in a natural environment, but similar large, saddle-shaped objects are rare. The observers in the current study, therefore, might have had less visual experience with these types of (i.e., saddle-shaped) background surfaces, and this lack of experience could have negatively affected their performance.

The current results show that the observers' recognition performances from deforming shadows were higher than those obtained from static shadows in most instances, regardless of the type of background surface (see Figures 4-8). Therefore, deforming shadows are more informative for the perception and recognition of 3-D object shape, even when the shadows were distorted by curved background surfaces. This result agrees with the findings of Norman et al. (2000) who found similar results for shadows cast onto flat background surfaces. The results also showed that it is sometimes difficult to recognize the 3-D shape of an object from static 2-D silhouettes, especially from static shadows projected onto the saddle. However, it is important to note that the observers' performance, even in the static conditions, was almost always well above chance (which was 20 percent correct). Therefore, even the static shadows projected onto the curved background surfaces contained large amounts of perceptually useful information.

The current experiment employed a set of smoothly curved 3-D objects and their 2-D cast shadows also had smoothly curved boundaries. Wallach and O'Connell (1953) believed from the results of their experiments that the deformations of smoothly curved boundary contours could not support the perception of a rigid object rotating in depth.

The results of the current experiment demonstrate that Wallach and O'Connell were not correct as regards to this point. In this regard, the results of this experiment confirm those of the earlier studies of Norman and Todd (1994), Norman et al. (2000), and Norman and Raines (2002). All of these studies, taken together with the current results, show that the deformations of smoothly curved shadows with no identifiable features, such as sharp corners, are sufficient for the recognition of 3-D object shape.

In summary, the results of the current experiment show that deforming shadows are informative for the perception of 3-D object shape and also support the idea that boundary contours contain a wealth of information to help human observers recognize naturally shaped objects. The findings of the current experiment also show that the perception and recognition of 3-D object shape is largely invariant over changes in 2-D projected shadows that are caused by the presence of convex curved background surfaces. Therefore, this experiment confirms and extends previous studies and leads to a better understanding of how human observers perceive and recognize 3-D shape from cast shadows in real environments.

References

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review, 61,* 183-193.

Koendrink, J. J. (1984). What does the occluding contour tell us about solid shape? *Perception, 13,* 321-330.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology, volume 1* (pp. 103-189). New York: John Wiley.

Miles, W. R. (1931). Movement interpretations of the silhouette of a revolving fan. *American Journal of Psychology, 43,* 392-405.

Norman, J. F., Dawson, T. E., & Raines, S. R. (2000). The perception and recognition of natural object shape from deforming and static shadows. *Perception, 29,* 135-148.

Norman, J. F., Phillips, F., & Ross, H. E. (2001). Information concentration along the boundary contours of naturally shaped solid objects. *Perception, 30,* 1285-1294.

Norman, J. F., & Raines, S. R. (2002). The perception and discrimination of local 3-D surface structure from deforming and disparate boundary contours. *Perception & Psychophysics, 64,* 1145-1159.

Norman, J. F., & Todd, J. T. (1994). Perception of rigid motion in depth from the optical deformations of shadows and occlusion boundaries. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 343-356.

Richards, W. A., Koenderink, J. J., & Hoffman, D. D. (1987). Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of the Optical Society of America A, 4,* 1168-1176.

Tse, P. U. (2002). A contour propagation approach to surface filling-in and volume formation. *Psychological Review, 109,* 91-115.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45,* 205-217.